DECISION SUPPORT SYSTEMS

FOR LEGAL RESEARCH AND ANALYSIS*

by

Jeffrey A. Meldman

CISR #38

Sloan School WP #1039-79

*A draft version of this paper was presented at the Second Annual Symposium on Computational Systems, Instituto Tecnologico y de Estudios Superiores de Monterrey, Monterrey, Mexico, 15-18 March 1977.
DECISION SUPPORT SYSTEMS FOR LEGAL RESEARCH AND ANALYSIS

Jeffrey A. Meldman
Massachusetts Institute of Technology
January 1979

The use of computers in the legal profession lags far behind computer use in the general business and professional community. The reasons for this are not hard to understand. Lawyers traditionally have been reluctant to jump too quickly into new forms of business technology. It has been estimated, for example, that lawyers took 20 years longer than the general business community to start to use typewriters. Whereas most managers and doctors, and other professionals today take pride in being as modern and scientific as possible, many lawyers are still suspicious (and often ignorant) of scientific methodologies. Sometimes these seem incompatible with long-standing (and often effective) legal methodologies.
Nevertheless, many lawyers have begun to use computers. As might be expected, some of the early applications included law office accounting, court administration, tax return preparation, title searching, and so forth. The computer is beginning to be used, however, to support the lawyer's tasks of legal analysis and research. It is the nature of these tasks and of this kind of decision support that we will now explore.

Consider a client consulting an attorney. The attorney may be asked for advice about the legal consequences of some contemplated activity. The attorney may have to draft a contract to protect the client's interests over some range of possible future situations. Or perhaps the client already has become involved in a predicament. He or she may wish to bring legal action to recover for losses or injuries caused by another, or wish to be defended in such a lawsuit-- or in a criminal proceeding.

In providing counsel in such situations, the attorney uses many forms of reasoning and skill. One of these is a fundamental technique that we will call legal analysis. By this term, we mean: the logical derivation of a legal conclusion from a particular factual situation in the light of some body of legal doctrine.

Before we can understand the nature of this process, we must be careful to distinguish it from the larger process of legal reasoning. This term is used generally to describe the process by which
judges decide cases (and, therefore, it is a part of the attorney's overall reasoning, too). In his classic treatise on the judicial process, Benjamin Cardozo separates four major paths along which the force of legal reasoning exerts itself. These are: logical progression, historical development, custom, and social justice. Regarding the force of logic, he explains:

In putting it first, I do not mean to rate it most important. On the contrary, it is often sacrificed to others. I have put it first because it has, I think, a certain presumption in its favor. . . . Lacking [some consideration of history or custom or justice], I must be logical.

* * *

You may call the process one of analogy or of logic or of philosophy as you please. Its essence in any event is the derivation of a consequence from a rule or a principle or a precedent which, accepted as a datum, contains implicitly within itself the germ of the conclusion.

What we are calling legal analysis corresponds substantially only to this first, logical component of legal reasoning. Legal analysis is performed on a particular set of facts against a background of legal doctrine. Cardozo is referring to such doctrine when he speaks of "a rule or a principle or a precedent."

Legal doctrine is embodied in different forms, such as in the statutes and the administrative regulations of our state and federal governments. In common-law systems such as that of the United States, the characteristic embodiments of legal doctrine are the judicial decisions in the individual cases. The doctrines of prior cases, which are called "the common law," serve as
precedents for future, similar cases, as we will discuss presently.

In common-law countries, cases are the most important form of legal
d Doctrine. Substantial areas of legal doctrine (especially of state
law) exist only in cases -- there are no statutory or regulatory
provisions that pertain. Even in areas of law covered by statutes
or regulations, these are always subject to the courts' interpretation
in individual cases, and these interpretations are binding as
precedents for future, similar cases.

We might state the rule of precedent like this:

The legal consequence of a particular factual situation (in a
given jurisdiction) must be the same as was the result of any
previous case (in that jurisdiction) that involved the same
factual situation.

If this doctrine is interpreted in its strictest possible sense,
it is virtually useless because of the small likelihood that the
particular set of facts at hand is going to be exactly "the same"
as the facts in some prior case. If the rule is to be meaningful, the
word "same" must be understood in the sense of "same kind of" or
"similar." Certain differences between the facts at hand the the
facts in the prior case must be ignored. Which ones? and how
large may the differences be? Having to answer these questions is
what makes the application of so simple-looking a rule so difficult.

In practice, this problem is approached through two mechanisms,
both involving generalization. In the first place, the holding of
a case (that is, the particular piece of doctrine for which the case
stands as authority) is almost always intended and understood in terms
more general than the specific persons, objects, activities, et cetera,
that were actually involved in the case. Judges understand well the
role that their written opinions play as precedents for future cases,
and they tend to write in terms of categories rather than individualities.

Consider the following holding from a 1956 North Carolina case involving the extension of surgery:

In short, where an internal operation is indicated, a surgeon may lawfully perform, and it is his duty to perform, such operation as good surgery demands, even when it means an extension of the operation further than was originally contemplated. Notice that the judge speaks not in terms of the individual defendant, but in terms of "surgeons," a category of persons that includes the defendant in this case as well as other surgeons who might be future defendants in similar cases. And the holding is not intended to be restricted to operations exactly like the one in this case, nor even to operations of a particular kind (this had been an appendectomy), but to the entire category of "internal operations."

What is the appropriate scope of such categories? This is determined by the reasoning that underlies the decision. In this case, Chief Justice Barnhill's stated reasoning included such factors as the known difficulty of presurgical diagnosis of internal ailments, the unavailability of obtaining further consent due to general anesthesia, and the desire to encourage surgeons who may be tempted to shirk from duty for fear of a lawsuit. (Remember that the process of judicial reasoning includes many other intellectual activities besides what we are calling logical analysis.) Clearly if this reasoning is valid for this individual defendant performing this particular operation, it is also valid for any surgeon performing any internal operation. Thus, he stated his holding in these
Ultimately, the holding in any case is determined not by the words of the judge who wrote the decision, but by future judges who interpret the decision. When a particular holding is thought to be unreasonably broad, which sometimes happens, it is interpreted more narrowly.

The second mechanism for generalization is analogy. This method is invoked when the facts at hand fall near, but not within, the scope of a prior holding. Once again the underlying reasoning of the prior case is the key. If that reasoning appears to apply with equal force to the facts at hand, then (absent other precedent, of course) the result in the prior case is controlling. Note that this is not a reassessment of the scope of the prior holding. It is an argument based on similarity, not on inclusion.

For example, consider the situation in which a dentist, while extracting a tooth from a patient under general anesthesia, discovers the necessity of extracting a second tooth, and does so without obtaining additional consent. These facts fall outside the categories "surgeon" and "internal operation" used in Barnhill's opinion. Yet it can be argued that much of Barnhill's reasoning applies equally well to these facts. On the other hand, consider the case of a garage mechanic who performs additional automobile repairs without the consent of the customer. Barnhill's reasoning is mostly irrelevant here, and the logic of analogy fails.
Through repeated use of analogy, the reasoning in individual cases gradually becomes extended to categories much broader than can be dictated by the facts of any single case. At some point, a perceptive judge may become aware of this growth, and reformulate a doctrine, or combine several doctrines, in appropriately broader terms.

When categories of fact become more general, it becomes more difficult to recognize whether individual facts do or not fit into the categories. The solution to this problem sometimes lies within the purview of legal analysis. There might, for example, be previous similar cases in which the judge ruled one way or the other at this point. Often, however, such questions are left to the intuitive reasoning of the fact-finder (e.g., the jury) in a trial.

Legal analysis applied to statutes and to regulations operates much in the same manner. One of the ways in which courts "interpret" statutes is to make decisions about which facts do, and which facts do not, fit into stated categories.

There are other ways in which doctrines interact with each other. It is very common, for example, for the facts presented by the client to involve several legal issues in a combination for which there is no comprehensive precedent. The issues must therefore be decomposed, and each analyzed with respect to a separate set
of precedents. Unfortunately, the issues in a case are not always independent, and the reasoning of an opinion may fail when the issues are treated separately.

It is also commonplace for the facts at hand to fall within, or equally close to, two apparently contradictory precedents. An attorney must often be able to argue that the reasoning of one authority applies more strongly to the facts at hand than does the reasoning of an opposing authority.

These are some of the complications of the task of legal analysis. When we add the components that Cardozo called history, custom, and social justice, as well as pragmatic considerations (like the disposition of a particular judge, or the likelihood of a cash settlement), we begin to see how complex the attorney's routine decision process can become. Hopefully, however, we can identify certain aspects of the process that are highly mechanical and involve little intuitive judgement. If so, a certain amount of computer support may be effective. We will return to this question later.

Let us first explore the process of legal research. This can be thought of as the search for the legal authorities -- statutes, regulations, and cases -- that pertain to the facts presented by the client. Very often, an attorney is able to proceed with some or all of the legal analysis without the need for formal research. He or she may already be familiar with the relevant authority for cases of this kind, or may be led to them directly by a law partner or fellow attorney who has handled such cases before. When such ad hoc
methods are not available, the attorney must spend some time in the library.

In the traditional research arrangement, the attorney finds certain basic kinds of resources in the library. First there are treatises, legal encyclopedias, and law reviews. These are called "secondary authority" because they contain restatements of legal doctrine composed by legal scholars. They do not have the weight of the doctrines contained in the "primary authority" (the statutes, regulations, and cases themselves). Treatises and encyclopedias are often used in the early stages of research to educate the researcher with respect to the particular issues involved in a particular area of law. Law Review articles provide similar instruction, but of very narrow issues.

Statutes both for the states and the Federal "U.S. Code" are relatively well organized. Most legislators attempt to structure their statutes so that doctrines pertaining to related areas of activity are grouped together in the same or adjoining segments. These chapters and sections are also given descriptive titles, and their contents are usually outlined at the beginning of each major segment. There is also often and index of words and phrases. These aids help considerably in the finding of relevant doctrine. Even so, the scheme of organization is not always sensible from the researcher's point of view, nor is it consistently adhered to. It is usually necessary to explore numerous approaches in order to find a section or sections that might apply to a particular set of facts.
Administrative regulations are less well organized. Federal agencies are required to publish their regulations. Some of these regulations (like those of the Internal Revenue Service) are organized directly in conjunction with related legislation (the U.S. Tax Code), but most are not. State regulations are usually more poorly organized and often available only on request from the agency or state capital. Sets of regulations may or may not have tables of contents or indexes. Because regulations are arranged separately by each agency, it is often difficult to find out if regulations pertaining to a particular subject matter even exist.

For case law, if it were not for separate means of assistance, the task of research would be hopeless. The decisions of the Massachusetts Supreme Judicial Court, for example, fill over 350 volumes, each of which contains an average of over 200 cases. Like the cases of all jurisdictions, these are printed and bound in simple chronological order. The titles of the cases indicate nothing but the names of the parties. There is no way in which they could be read or even sampled in search of relevant doctrine.

Indirect methods of case-law research had to be developed long before the advent of computer technology. We already mentioned the restatements of legal doctrine found in legal encyclopedias, treatises, and law review articles. These sources of "secondary authority" are produced and organized in terms of major legal topics
(like the law of torts), subtopics (battery, assault), and so on.
Besides providing the lawyer with smaller, compartmentalized bodies of doctrine with which to perform preliminary analysis, these materials usually contain references to the cases on which the restatements are based. Such cases often turn out not to be directly relevant, but they are a valuable entry point into the case law itself.

Other means specifically intended to assist case research have been devised. The West Publishing Company has promulgated a widely used key-number system. The system is based on a hierarchical arrangement of legal issues that runs from major topics at the top all the way to specific holdings of cases at the bottom. Numerical values are assigned to these specific issues such that (theoretically) similar holdings receive the same, or adjacent, numbers. West publishes a digest of case law in which can be found the restated holdings of cases, arranged according to these numbers. West's encyclopedia, Corpus Juris Secundum, is also arranged according to these numbers, and the reprints of cases that West publishes for every jurisdiction have these key numbers inserted appropriately in the text. Another publisher, Shepard, produces a citation index that lists, for each case within a given jurisdiction, all later cases in which the judge makes some reference to that particular case. By "shepardizing" a case, a researcher can determine where the holding of that case had played a role in the reasoning of more recent decisions. There are also annotated versions of state and federal statutes that contain digested holdings of cases that relate to, or interpret, the individual sections of the statutes.
To a large extent, the process of legal analysis involves standard problems of document retrieval. The secondary authorities, the tables of contents, the indexes, the digests, all are systems for identifying cases and sections of statutes and regulations potentially relevant to the facts under analysis. Ultimately, these documents must be physically retrieved from the library shelves and examined by the researcher. We might represent this process graphically as is shown in Figure 1. The attorney begins with a set of facts. It may or may not then be necessary to consult a treatise or encyclopedia in order to identify the specific legal issues that are raised and the bodies of law (such as the regulations of particular agencies) that might contain relevant doctrine. The separate identification systems for these bodies of law can then be used. Each system requires the user to supply a certain amount of "input." For example, in using an index, the words the user attempts to look up are the input. The headings in a table of contents and in a digest system operate in much the same way. In traditional manual systems, these input terms are based mainly on legal issues (such as "negligence"), and to a lesser extent on factual components (such as "automobiles").

Inputs may or may not yield "outputs," that is citations of potentially useful authorities. Often the inputs will be dead ends; sometimes they will lead to suggestions by the system for other input
FIGURE 1: A Document-Retrieval Model of Legal Research
attempts, as in the case of cross references. Some identification systems permit input browsing. Most tables of contents, for example, and digest systems like West's, allow for a certain amount of browsing because similar subject matter tends to be arranged adjacently.

Some of the output citations suggested by the identification systems can be eliminated by the researcher when they appear to be irrelevant. Digested holdings are particularly useful in this process, although many of West's digests are too brief to reveal the relevancy of the digested authority. Outputs that are not eliminated at this stage are then retrieved and examined. These may then be rejected as useless or retained for further use.

If we examine this process from the point of view of the decisions that the researcher is required to make, we might identify them as follows:

(1) What legal issues are raised by the facts?
(2) Which bodies of law are likely to contain doctrines relevant to these issues and facts?
(3) Which identification systems for these bodies of law should be tried?
(4) What inputs to these systems should be used?
(5) Given the suggestions by a system for other inputs, which should be tried?
(6) Given the outputs of a system, which authorities should be physically retrieved and examined?
(7) Given a retrieved authority, is it of any use?
If we were now to integrate this conceptualization of legal research with some of the analysis tasks we discussed earlier, we might produce an overall model of legal analysis as shown in Figure 2. This figure points out several important aspects of the relationship between research and analysis. Notice first that the retrieval task is sandwiched between two forms of analysis. During the preliminary analysis, the facts elicited from the client, together with the attorney's general knowledge of law suggest the legal issues and the bodies of law with which the attorney begins his or her research. This preliminary analysis also suggests further questions of fact to be asked of the client. It is this stage of analysis in which the attorney may be assisted by secondary legal authority.

The other form of analysis is performed as each step of the research is completed, that is, as each piece of useful authority is examined. The major aspect of this analysis, as we described it earlier, is the determination of the extent to which the facts at hand fall within, or close to, the doctrine stated in the authority. Because of the complexities of this analysis, examination of each piece of authority may also suggest to the attorney new leads for research -- new issues, new bodies of law, new questions of fact, other relevant authorities. These suggestions easily lead back to the earlier stages of analysis and research, which results in a highly iterative process.
FIGURE 2: An Integrated Model of Legal Analysis and Research
The analysis components of this process also can be examined in terms of the decisions that the attorney is required to make. In the preliminary analysis we can identify decisions such as:

(A) What issues are suggested by the facts elicited so far?
(B) What further facts are needed?
(C) What secondary authorities might be helpful in this analysis?

Eventually this leads to decisions 1 and 2, stated above, regarding the legal issues and bodies of law to be used for research.

The authority analysis involves decisions such as:

(V) How well do the facts at hand fall within, or analogously close to, the scope of the doctrine expressed in this authority?
(W) How well does the reasoning expressed in this authority apply to the facts at hand?
(X) Does this authority suggest any other potentially relevant authorities?
(Y) Does this authority suggest new legal issues, new bodies of law, or new questions of fact?

Finally, as each iteration proceeds, it is necessary to decide:

(Z) Is further research indicated?

With this decision-oriented model in mind, let us briefly examine three typical computer-based research aids that have been made available to the legal community.

The first research systems that took effective advantage of the computer's high processing speed were the full-text index, logical inquiry systems, sometimes called "key-word-in-combination." Their use in law was developed by John Horty, who was himself engaged in legal research involving the health statutes of the various states. The differing methods by which the states organized their statutes
was presenting severe problems -- especially for those states where laws related to health were scattered and buried in laws pertaining to other matters. Horty's solution was to load the full text of a statute into a computer file, have the computer remove the insignificant words (like a, the, and), and then have the computer produce an index listing the exact locations in the statute of each of the significant words. With this index in the machine's memory, he could ask the computer to identify or to print out every statutory section containing a particular word, like "health." Or he could specify a logical combination of words, and ask, for example, for all sections containing both "health" and either "dangerous" or "hazardous." By 1970 this technique was applied to case law as well as to statutory law. In addition to simple logical word combinations, these systems can find word coincidences within given intervals, search for families of words with the same roots, and sometimes, generate and search for synonyms. Queries usually can be modified interactively, so that constraints can be added or relaxed. The systems also permit on-line inspection of entire documents or, if desired, of just those portions of the documents where the key words were located. These systems are available commercially but are being used by a relatively small number of relatively large law firms.

In terms of our model, the most significant feature of these systems is the way they alter the nature of the input used for document identification (decision 4). In the manual systems, a researcher usually must try input terms based on the way a cataloger has characterized the legal issues in the case. In the logical inquiry systems, the researcher must anticipate unique combinations of
terms -- factual as well as legal -- likely to be used by a judge in deciding a relevant case. While there are those who debate as to which of these input schemes is more powerful, a system that permitted both forms of input would probably be better than either alone.\textsuperscript{11}

Perhaps the most attractive feature of the full-text systems is the tight coupling between document identification, physical retrieval, and partial (or full) examination. This makes it significantly easier and faster to determine if a document is likely to be useful or not (decisions 6 and 7). Finally, these systems may be helpful in supporting pragmatic decisions that lie outside our model. Some lawyers, for example, use them to study the opinions of particular judges. (These are collectively retrievable by using the judge's name as an input term.)

A second type of computerized case research system, developed in the mid-sixties, was the citation-linkage system.\textsuperscript{12} The user of this system had to supply as input the citations of two cases already known to be relevant to the facts under analysis. The system then generated forward and backward citation links for that case (cases cited in the case and cases citing the case). This linkage was then repeated for the generated citations until a certain cut-off level was reached. The resulting network of citations was searched for cases with certain frequencies of appearance, and the citations of these cases were listed as output.
This was basically a document identification system, the input for which required prior manual research for (or knowledge of) relevant cases. It performed its own very simple document analysis by which other potentially useful documents were identified (question X). However, the system did not include means for text retrieval for examination by the attorney, and it was operated in batch mode. This made iteration back to the input stage very cumbersome. The system was not commercially successful and is no longer available.

The newest system, just now coming into commercial operation, is a logical inquiry system based on the "full" text of the West digest entries. The texts of the cases themselves are not available for query or for examination. The system amplifies the usefulness of the digests by providing iterative interaction in the finding of promising citations (decisions 4 and 5). This can result in a greater number of outputs, and it can make it easier to determine whether a given output is worth retrieving for examination (decision 6). Because the digests are quite brief and because the same words tend to be used with high frequencies, this system tends to generate a large number of not very relevant cases. The current attempt to correct for this problem is to display the output digests in order of their coincidence frequencies.

While we have been able to identify some of the relationship between these systems and the decisions enumerated in our model of legal analysis, it is clear that they were not really designed from a decision-support point of view. In many cases, they were developed in a manner sometimes called "have solution, will travel," that is, from techniques already used in other document retrieval systems, or
from existing (and proprietary) data bases. In particular, systems
like these tend not to address the process of analysis as a whole. Thus,
there has been little concern with support for the decisions involved
in the preliminary analysis stage (decisions A,B,C,1, and 2) or
in document analysis (decision V,W,X,Y and Z).

The analysis decisions, of course, are more difficult, and they
are more richly impregnated with the need for human judgment. There
is, however, considerable room even here for mechanical assistance
that can help direct an attorney's analysis in potentially fruitful
directions. There has been some research along these lines, although
no such systems are yet available. There is one technique, for
example, in which a machine representation is built of the boolean
logic relating legal issues to their sub-issues. This system
traces through the logic, asking the attorney's opinion as to the
resolution of each issue, thereby identifying those issues for which
additional facts or additional research are required. Another
technique uses machine representations of the factual patterns
contained in the legal doctrines of both secondary and selected
primary authorities. The attorney describes the factual situation
to be analyzed in a similar format. The system tries to match the
factual patterns so as to identify the legal issues that may be
involved, and where possible to identify potentially relevant authorities.
The analysis of these authorities, particularly with regard to the
reasoning in the cases, is still left for the attorney's judgment.
Nonetheless, this system can bring the attorney directly from the
stage of preliminary analysis to the analysis of the authorities without
necessitating some of the decisions involved in using document identifi-
cation systems.
The future of research efforts of this kind is not entirely clear. Many of the problems that will need to be addressed are sociological as well as technical. The growing use of computer technology in the law brings to mind, for example, an uncomfortable spectre of an overly mechanical system of justice. Hopefully, this possibility will be avoided. Computer designers and computer users in many other fields are coming to understand the importance of separating clearly the role of machine assistance from that of human judgment. This has been the dominant theme in the development of—and even the name—"Decision Support Systems." The theme is based as much on the goal of system effectiveness as on considerations of humanism. As long as the role of human legal judgment is understood in its proper relation to the more mechanical aspects of legal process, decision support systems may be able to bring the legal community safely into the computer era.
REFERENCES


7. Cardozo, pp. 31, 49.


12. This system was developed by the Data Retrieval Corporation of America, Milwaukee.

